View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Yale University

Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Public Health Theses

School of Public Health

January 2020

Use Of Shotgun Metagenomic Sequencing To Determine How The Human Gut Microbiome And Antibiotic Resistome Influence The Risk Of Recurrent Clostridioides Difficile Infection

Maria Alexandra Correa
mcorrea44@gmail.com

Follow this and additional works at: https://elischolar.library.yale.edu/ysphtdl

Recommended Citation

Correa, Maria Alexandra, "Use Of Shotgun Metagenomic Sequencing To Determine How The Human Gut Microbiome And Antibiotic Resistome Influence The Risk Of Recurrent Clostridioides Difficile Infection" (2020). *Public Health Theses*. 1930.
https://elischolar.library.yale.edu/ysphtdl/1930

This Open Access Thesis is brought to you for free and open access by the School of Public Health at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Public Health Theses by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Use of shotgun metagenomic sequencing to determine how the human gut microbiome and antibiotic resistome influence the risk of recurrent *Clostridioides difficile* infection

Maria A. Correa

Year Completed: 2020
Year Degree Awarded: 2020
Master of Public Health
Yale School of Public Health
Primary Adviser: Melinda Pettigrew, PhD
Secondary Adviser: Brian Weiss, PhD

Abstract

Clostridioides difficile is an opportunistic enteric pathogen that can cause a range of symptoms in humans from diarrhea to pseudomembranous colitis. The risk of recurrence of *C. difficile* infection after initial treatment with antibiotics is high. Although prior studies have sought to understand the link between the human gut microbiome and the risk of recurrence, none have utilized shotgun metagenomic sequencing methods to establish a relationship between microbiota diversity and recurrence risk. In this study, stool samples were obtained from 47 patients at Yale New Haven Hospital who tested positive for an incident *C. difficile* infection. Shotgun metagenomic sequencing was used to characterize the abundance and diversity of the microbiota and the antibiotic resistome of each sample. The association between taxonomic diversity of the gastrointestinal microbiota and the risk for recurrence, case definition, and patient characteristics was analyzed. There was a significant association between age and the risk for recurrent infection, with older patients more likely to experience recurrence. Linear discriminant analysis effect size revealed that certain taxonomic groups were differentially prevalent in patients with and without a recurrent episode and in patients with community-acquired infections vs. those with hospital- or healthcare-associated infections. Those who experienced a recurrent infection had a higher abundance of *Blautia producta*, and a lower of abundance of *Gardnerella vaginalis* and *Eggerthella lenta* than patients who did not experience a recurrence. The microbiota of patients who had a community-acquired infection were more likely to include protective bacterial species and bacteriophage, such as *Bacteroides*, *Faecalibacterium prausnitzii*, and crAssphage. This study provides a preliminary exploration into the association between the human microbiome and the risk of *C. difficile* recurrence and demonstrates the value of using shotgun metagenomic sequencing for further investigation.

Acknowledgments

I would like to thank my thesis advisers Melinda Pettigrew and Brian Weiss for their help and support throughout the research process. I would also like to thank Danyel Olson for help with the patient chart reviews and Martina Wade for providing help in the lab.

Table of Contents

Tables and Figures

Table 1. Description of the sample, by CDI recurrence.[a]

Characteristic	Recurrence		p[b]
	Yes (N = 7)	No (N = 35)	
Age (years)	75.1 ± 12.9	54.9 ± 24.1	0.03
Sex			
Male	3 (42.9)	17 (48.6)	1.00
Female	4 (57.1)	18 (51.4)	
Case Definition			0.73
Hospital-acquired	2 (28.6)	11 (31.4)	
Healthcare- associated	5 (71.4)	18 (51.4)	
Community-acquired	0 (0.0)	5 (17.1)	
Chronic Kidney Disease			0.16
Yes	3 (42.9)	6 (17.1)	
No	4 (57.1)	29 (82.7)	
Congestive Heart Failure			0.19
Yes	2 (28.6)	3 (8.6)	
No	5 (71.4)	32 (91.4)	
Stroke			0.53
Yes	1 (14.3)	3 (8.6)	
No	6 (85.7)	32 (91.4)	
Diabetes mellitus			0.11
Yes	3 (42.9)	5 (14.3)	
No	4 (57.1)	30 (85.7)	
Antibiotic use ≤ 12 weeks before incident CDI			1.00
Yes	6 (85.7)	28 (80.0)	
No	1 (14.3)	5 (14.3)	
Unknown	0 (0.0)	2 (5.7)	
H2 blocker use ≤ 12 weeks before incident CDI			1.00
Yes	3 (42.9)	17 (48.6)	
No	4 (57.1)	17 (48.6)	
Unknown	0 (0.0)	1 (2.9)	
Proton pump inhibitor use ≤ 12 weeks before incident CDI			0.53
Yes	2 (28.6)	16 (45.7)	
No	5 (71.4)	18 (51.4)	
Unknown	0 (0.0)	1 (2.9)	

[a] Table values are mean ± SD for continuous variable and n (column %) for categorical variables.
[b] P-value is for Duncan's MRT (continuous variable) or Fisher's Exact test (categorical variables).

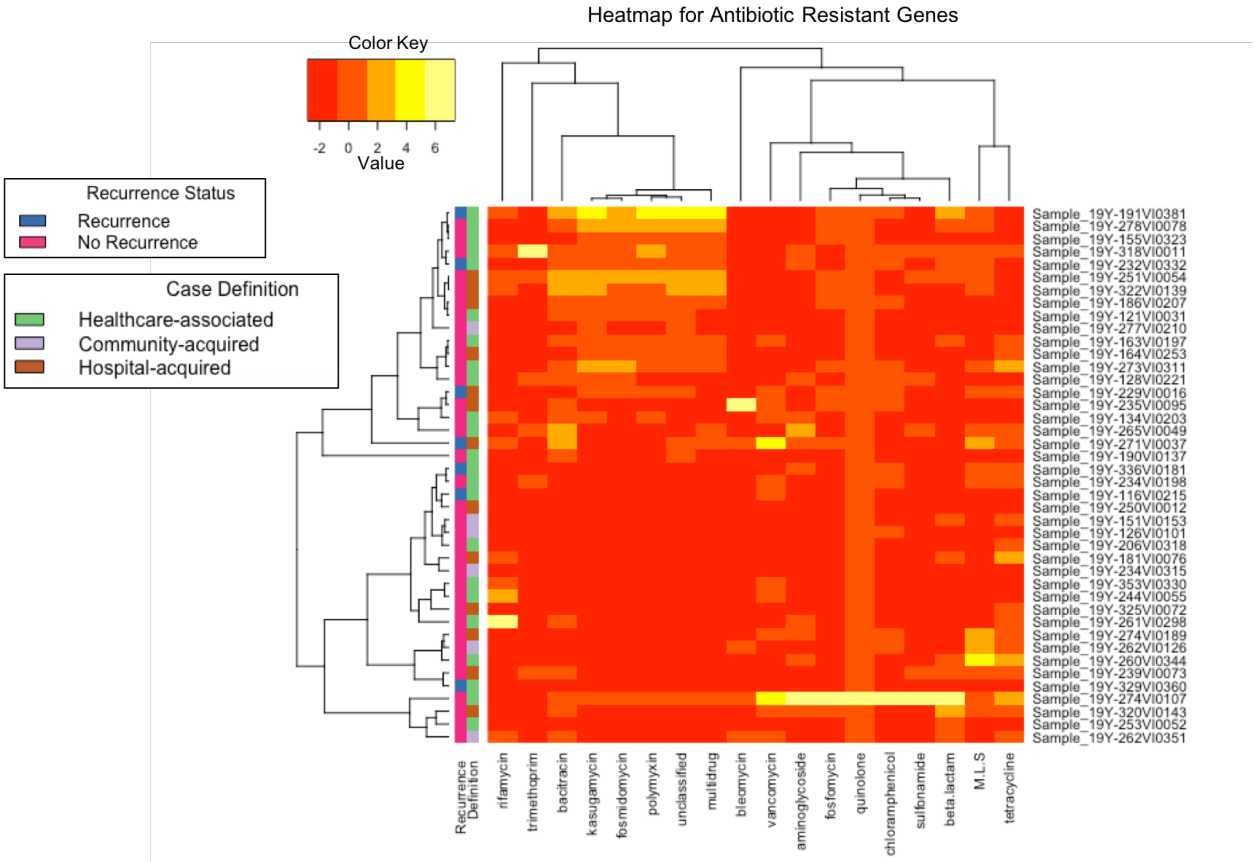

Figure 1. Heatmap of the most abundant antibiotic resistance genes in stool samples from 42 patients. The color keys for recurrence status and case definition are indicated on the left. Abbreviation: MLS, macrolide-lincosamide-streptogramin.

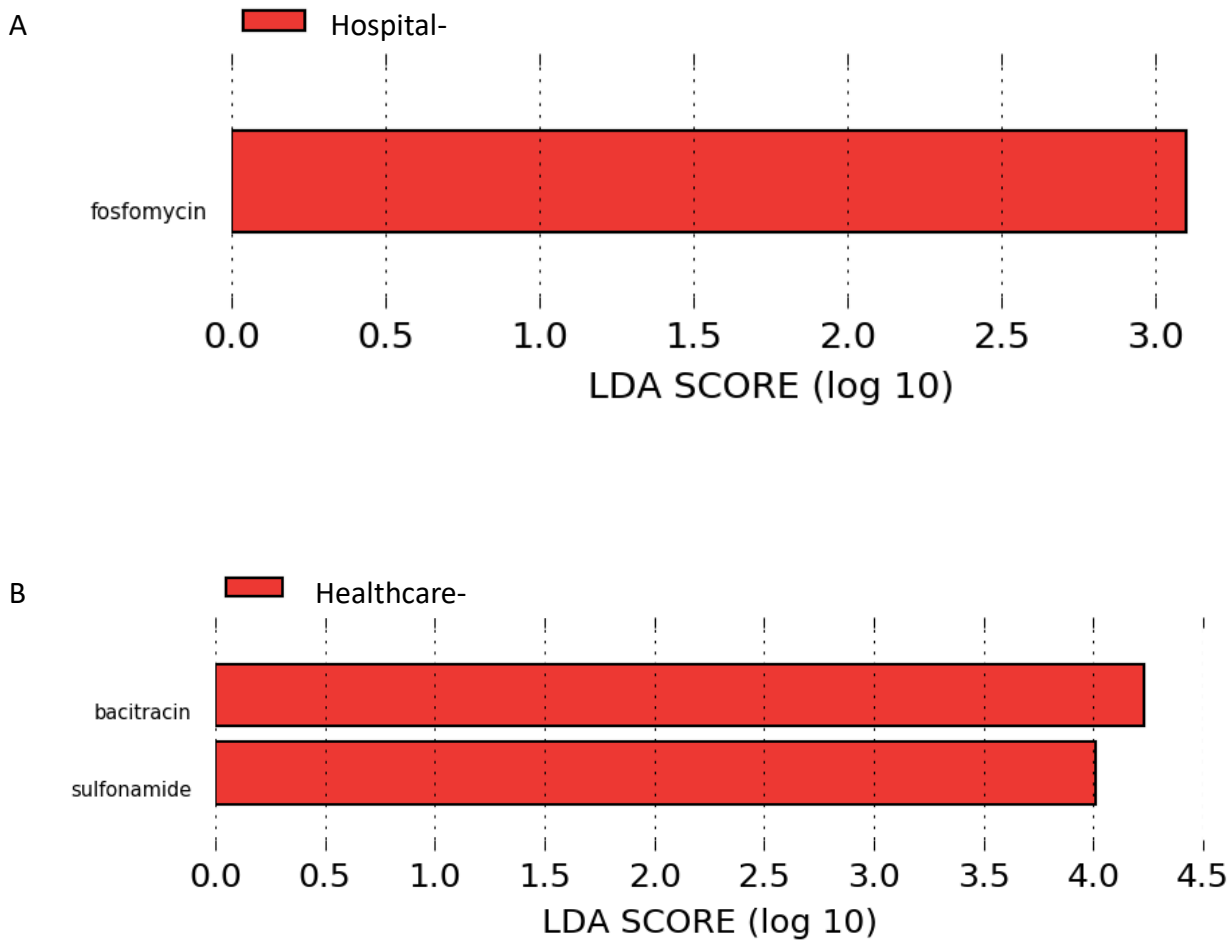

Figure 2. Linear discriminant analysis effect size scores for comparisons between relative abundance of antibiotic resistance genes among 47 patients based on case definition. *A*. Hospital-acquired infections vs. community-acquired infections. *B*. Healthcare-associated infections vs. community-acquired infections. Abbreviation: LDA, linear discriminant analysis.

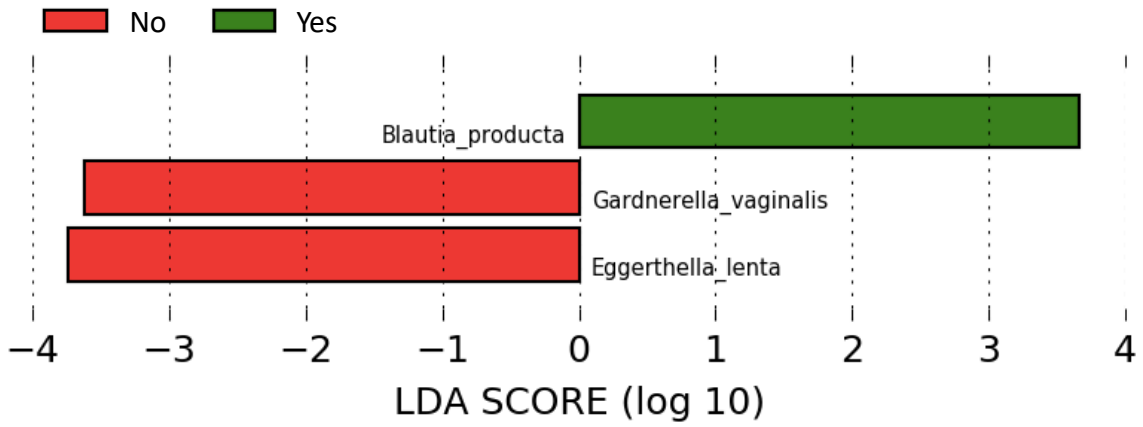

Figure 3. Linear discriminant analysis effect size scores for comparisons between relative abundance of taxa among 42 patients with or without a recurrent infection. Abbreviation: LDA, linear discriminant analysis.

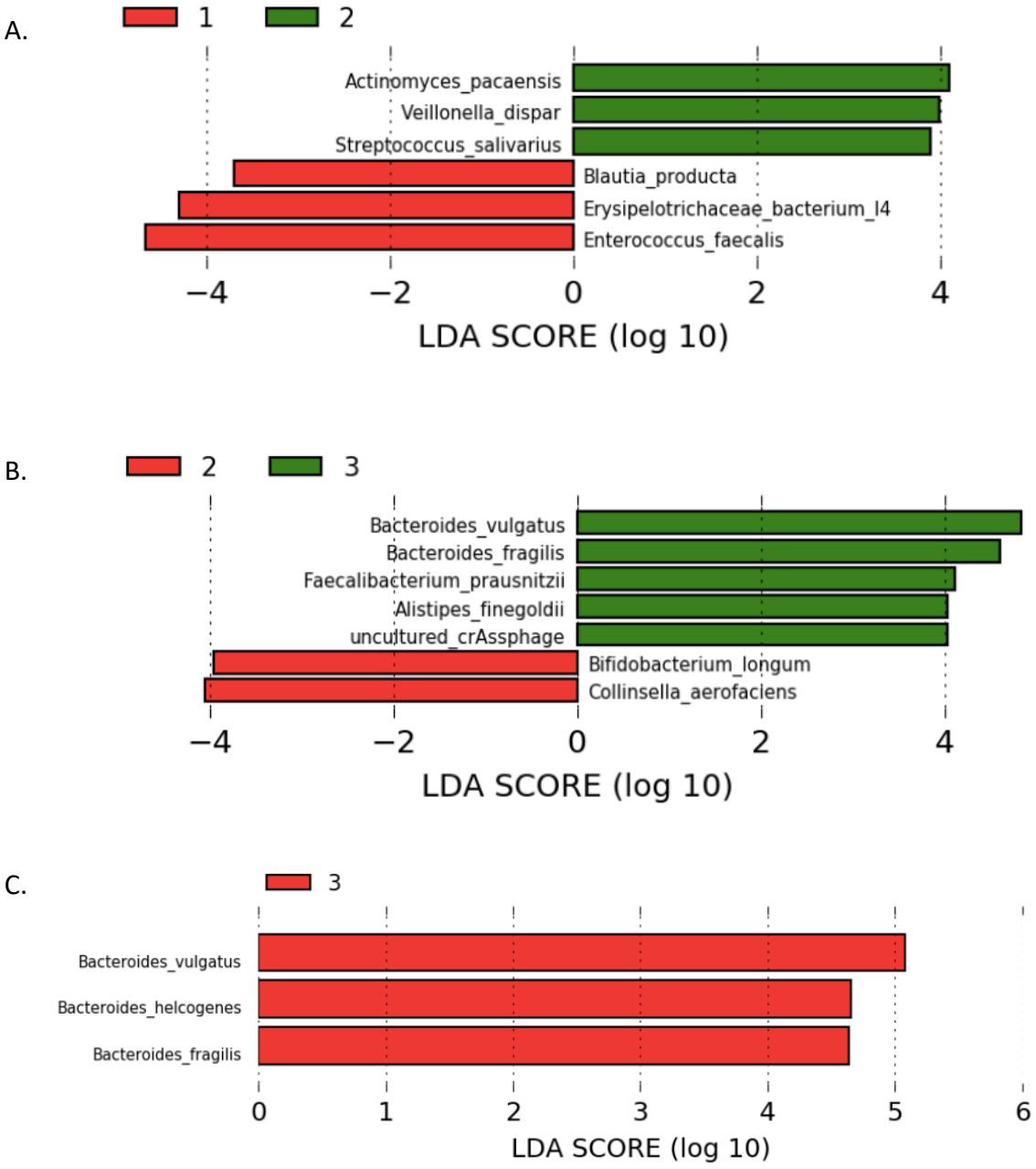

Figure 4. Linear discriminant analysis effect size scores for comparisons between relative abundance of taxa among 47 patients based on case definition. *A.* Hospital-acquired infections (1) vs. healthcare-associated infections (2). *B.* Hospital-acquired infections (1) vs. community-acquired infections (3). *C.* Healthcare-associated infections (2, not shown) vs. community-acquired infections (3). Abbreviation: LDA, linear discriminant analysis.

Introduction

Clostridioides difficile is a gram-positive, spore-forming bacillus that causes diarrhea and colitis in susceptible individuals. It is a major cause of healthcare-associated infections in the United States, with upwards of half a million infections and nearly 30,000 associated deaths every year.[1] Up to 20% of all individuals who are infected with *C. difficile* experience a recurrence in infection within twelve weeks of treatment.[2] As such, *C. difficile* places a large burden on the US healthcare system, costing nearly $1.5 billion total annually.[3] Although initially considered to be a primarily nosocomial or healthcare-associated infection, community-acquired *C. difficile* infections (CDIs) have become an increasing problem, especially among populations previously thought to be at low risk for infection.[4]

C. difficile is resistant to a wide array of commonly used antibiotics, including penicillin, cephalosporins, clindamycin, tetracyclines, and fluoroquinolones.[5] Thus, it is an opportunistic pathogen that takes advantage of disturbances in the normal human gut microflora to proliferate and cause disease in human hosts.[6] One of the major risk factors for CDI, therefore, is repeated or sustained antibiotic usage.[7] Other known risk factors for infection are overnight stays in hospitals and long-term care facilities, advanced age, underlying medical problems, reduction of stomach acid through proton pump inhibitors, and the use of immunosuppressants.[8-10] The risk factors for recurrent infection are similar to that of incident infection and include increasing age and the severity of primary infection.[11]

Understanding the role of the gut microbiome in *C. difficile* infection is essential to understanding the risk factors for infection and developing targets for infection prevention and treatment. Certain bacterial species in the human gut microbiome have been shown to be associated with resistance to colonization by *C. difficile*, specifically members within

Ruminococcaceae, Lachnospiraceae, *Bacteroides*, and Porphyromonadaceae.[12] Infection with *C. difficile* has also been associated with a higher abundance and lower diversity of bacteriophage *Caudovirales*.[13] A metagenomic analysis of CDI-infected elderly patients in a hospital setting revealed that *C. difficile* infection was associated with a lower gut microbial diversity, as well as a lower incidence of commensal microbes and a higher incidence of opportunistic pathogens.[14] Furthermore, studies in infants suggest that a machine learning approach can be used to identify antibiotic resistance genes within the gut that can predict the effect of antibiotics on the gut microflora, a tool that could be utilized for predicting treatment outcomes for CDI.[15]

The association between the human gut microbiome and the risk of CDI recurrence has been explored using 16S rRNA gene sequencing. One such study found that patients who experienced a recurrent episode had elevated levels of *Veillonella*, Enterobacteriaceae, *Streptococcus*, *Parabacteroides* and Lachnospiraceae.[16] Although the link between the human microbiome and CDI risk has been investigated in previous studies, shotgun metagenomic sequencing has not yet been used to determine the relationship between the human gut microbiome and antibiotic resistome and the risk for CDI recurrence.

The goal of this study was to use shotgun metagenomic sequencing methods to determine the association between the risk of recurrence and the following factors: the abundance and diversity of specific bacterial taxa within the gut microbiome, and the presence of antibiotic resistance genes.

Methods

Yale University's institutional review board approved this study. Stool samples were obtained through collaboration with Connecticut's Emerging Infections Program, which is responsible for *C. difficile* surveillance within the New Haven catchment area. From April 26,

2019 to December 19, 2019, a total of 47 stool samples were collected from patients at Yale New Haven Hospital who tested positive for an incident *C. difficile* infection. A case was defined as incident if the individual was not treated for CDI in the twelve weeks prior to the date of stool collection. A medical chart review was conducted for each individual to identify possible risk factors for infection, including past medical history, antibiotic use, use of proton pump inhibitors, H2 blockers, and other immunosuppressants, as well as details on CDI treatment and recent contact with healthcare.

Individuals were classified based on recurrence status and case definition. A case was defined as recurrent if an individual experienced a second CDI within 2-12 weeks after incident infection. Any infection after this time period was considered a re-infection and not counted among the recurrent cases. Additionally, all CDIs included in the study were classified into one of three case definition groups: hospital-acquired infections, healthcare-associated community onset infections, and community-acquired infections. An infection was considered to be hospital-acquired if the incident stool sample was collected at least 3 days after admission to a hospital. In contrast, an infection was considered to be healthcare-associated community onset if the incident stool sample was collected less than 3 days after admission to a hospital and/or the patient had any of the following encounters with health care within 12 weeks of the date of the incident stool sample collection: visited a hospital emergency department, received surgery or dialysis treatment, or was admitted to a hospital. A patient's CDI was considered community-acquired if they did have any health care encounters within 12 weeks prior to the incident stool sample collection and/or their stool sample was collected less than 3 days after admission to a hospital.

Microbial DNA was extracted from stool samples using the PureLink Microbiome Purification Kit (Invitrogen). After DNA quantification using the Quant-iT Assay Kit

(Invitrogen), samples were prepared for sequencing using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina (New England Biolabs). Shotgun metagenomic sequencing was carried out on an Illumina HiSeq4000 at the Yale Center for Genome Analysis, after which taxonomic identification of bacteria and bacteriophage was carried out. Taxonomic labels were assigned to DNA sequences using Kraken, and species abundance was estimated using Bracken.[17,18] Antibiotic resistance genes (ARGs) were identified and quantified using ARGs-OAP v2.0.[19] ARG copy numbers were normalized against reference sequence lengths and then against prokaryotic cell number, and were grouped into the following ARG types based on mechanism of resistance: aminoglycoside, bacitracin, beta-lactam, bleomycin, chloramphenicol, fosfomycin, fosmidomycin, kasugamycin, macrolide-lincosamide-streptogramin, polymyxin, quinolone, rifamycin, sulfonamide, tetracycline, trimethoprim, vancomycin, and multidrug.

After the microbiome and antibiotic resistome for each sample were characterized, statistical methods were used to determine the unadjusted associations between case definition or the risk for recurrence and the following variables: age, sex at birth, underlying medical conditions, and the use of antibiotics, H2 blockers, and proton pump inhibitors prior to incident infection. Due to the low sample size, categorical variables were analyzed using Fisher's Exact test. Age was analyzed using Duncan's new multiple range test or one-way analysis of variance (ANOVA) as appropriate. All analyses were carried out using SAS v9.4.

Linear discriminant analysis effect size (LefSe) was used to identify differences in the presence and abundance of antibiotic resistance genes in patients who experienced a recurrent episode and those who did not. LefSe was also used to compare differences between the three case definitions by analyzing each pair of categories separately: community-acquired vs. hospital acquired, community-acquired vs. healthcare-associated, and hospital-acquired vs. healthcare-

associated. An alpha value of 0.1 and a threshold on the logarithmic linear discriminant analysis (LDA) score for discriminative features of 2.0 were used.

LefSe was also used to compare the taxonomic diversity between patients with and without a recurrent episode and between case definitions. Before analysis, the proportion of total sequencing reads was recalculated for each species in each sample after the removal of human DNA. The proportion of sequencing reads for each taxonomic group was summed across all samples, and the top 65 taxa with the highest total relative abundance were selected for analysis. The combined dataset was restricted to these 65 species while the rest of the taxonomic groups were collapsed into a single category for each sample. An alpha value of 0.1 and the and a threshold on the LDA score for discriminative features of 2.0 were used.

Results

A total of 47 stool samples were obtained for metagenomic analysis. Out of the 47 individuals who provided a stool sample, 25 were male and 22 were female, with ages ranging from 3 to 95. Of the total, 16 infections were classified as hospital-acquired, 25 were classified as healthcare-associated community onset, and 6 were community-acquired. Five of the individuals in the study died before the period during which a recurrent *C. difficile* episode may have occurred. These subjects were excluded from all analyses that used recurrence as a dependent variable. The resulting sample of 42 individuals included 20 males and 22 females. Of this sample, 7 individuals experienced a recurrent episode 2-12 weeks after incident infection.

The majority of individuals had underlying medical conditions prior to their incident *C. difficile* infections. The most common medical conditions were chronic kidney disease (n=11), diabetes mellitus (n=9), congestive heart failure (n=7), recent solid organ transplants (n=7), and hematologic malignancies (n=7). Frequency of medication use within 12 weeks of the incident

positive stool sample was also fairly high; 20 subjects took proton pump inhibitors, 21 took H2 blockers, while 29 took some other type of immunosuppressive therapy, such as steroids and/or chemotherapy. Antimicrobial therapy within the 12 weeks of incident stool collection was extremely common; 39 individuals were prescribed at least one antibiotic during that time period. The most common antibiotics were piperacillin/tazobactam (n=13), vancomycin (n=12), ciprofloxacin (n=11), ceftriaxone (n=11), and metronidazole (n=8). Cephalosporins were the most commonly administered class of antibiotics (n=21), followed by penicillins (n=20), and fluoroquinolones (n=14).

Age was found to be significantly associated with the risk for recurrence. Older people were more likely to experience a recurrent episode than younger people (p=0.038; Table 1). No other significant associations were observed between recurrence and other exposures of interest, including case definition, comorbidities, antibiotic use, and H2 blocker and proton pump inhibitor use. There were also no significant associations between these variables and case definition.

Hierarchical clustering was used to group samples by the abundance of antibiotic resistance genes (Figure 1). Samples did not cluster based on recurrence or case definition. However, LefSe indicated that genes that conferred resistance to fosfomycin were more abundant in patients who experienced hospital-acquired infections compared to those who experienced a community-acquired infection (Figure 2a). Additionally, bacitracin and sulfonamide resistance genes were more abundant in healthcare-associated community-onset infections than in community-acquired infections (Figure 2b). There were no significant differences in the abundance of ARGs between hospital-acquired and healthcare-associated

infections. Similarly, no differences were observed in the abundance of ARGs between patients who did and did not experience a recurrent infection.

A total of 6059 bacterial taxa and bacteriophage were identified in stool samples from 47 patients. LefSe identified certain taxa as differentially abundant depending on a patient's recurrence status. Patients who experienced a recurrent infection had a higher abundance of *Blautia producta* and a lower of abundance of *Gardnerella vaginalis* and *Eggerthella lenta* than patients who did not experience a recurrence (Figure 3). Differences in taxonomic abundance were also observed between case definitions. Compared to healthcare-associated infections, hospital-acquired infections had a higher abundance of *Actinomyces pacaensis*, *Veillonella dispar*, and *Streptococcus salivarius*, and lower abundance of *B. producta*, Erysipelotrichaceae, and *Enterococcus faecalis* (Figure 4a). In comparison to community-acquired infections, hospital-acquired infections had a higher abundance of *Bifidobacterium longum* and *Collinsella aerofasciens*, and a lower abundance of *Bacteroides vulgatus*, *Bacteroides fragilis*, *Faecalibacterium prausnitzii*, *Alistipes finegoldii*, and crAssphage (Figure 4b). Finally, patients with community-acquired infections had a higher abundance of *B. vulgatus*, *Bacteroides helcogenes*, and *B. fragilis* than patients with healthcare-associated infections (Figure 4c).

Discussion

The results from this study provide preliminary evidence that the risk of recurrence of *C. difficile* infection is associated with the diversity and composition of the human gut microbiome. Metagenomic analysis of stool samples from 42 patients suggests that *B. producta* is associated with the risk of recurrence. This finding is somewhat surprising because *Blautia* is a commensal genus that has been shown to be underrepresented in *C. difficile*-positive stool samples.[20] In one study, *B. producta* was used with a cocktail of nine other bacterial species to successfully treat

chronic CDI in five patients.[21] All of the patients in this study, however, were already colonized with *B. producta* prior to treatment, making it unlikely that *B. producta* alone was responsible for clearing the recurrent infections. Risk of recurrence was also negatively associated with *E. lenta*, which is part of the normal human gut microflora and can cause ulcerative colitis in humans.[22] Similarly to *C. difficile*, antibiotic treatment has been shown to increase the amount of *E. lenta* in the human gut microbiome.[22,23] The results from the current study suggest that the relationship between recurrent CDI and certain commensal microorganisms is more complex than originally assumed and warrants further investigation.

Bacterial taxonomic groups also differed in relative abundance between case definitions. A major result was that the microbiomes of patients with community-acquired infections were more likely to contain taxa that are associated with healthy human gut microflora. Community-acquired infections had a higher abundance of species in the genus *Bacteroides*, including *B. fragilis*, *B. vulgatus*, and *B. helcogenes*. *Bacteroides* species are commensal microorganisms that operate in the healthy human gastrointestinal tract.[24] One study showed that patients with CDI were deficient in *Bacteroides*, while healthy controls maintained relatively high levels of species from this genus. Compared to hospital-acquired infections, community-acquired infections were also associated with higher levels of *F. prausnitzii* and crAssphage. *F. prausnitzii* is a member of Ruminococcaceae and is the most abundant species found in a healthy human gut microbiome.[25] CrAssphage, a lytic phage that targets *Bacteroides*, is one of the most commonly found bacteriophages in the human gut.[26] Studies that sought to establish the association between bacteriophages and recurrent CDI risk and treatment found that individuals that suffer from recurrent CDI have lower levels of crAssphage than healthy fecal transplant donors.[27]

The abundance of healthy commensal bacteria and bacteriophage in patients with community-acquired infections may be due to a number of reasons. First, community-acquired infections are usually less severe than hospital-acquired infections,[4] which may indicate that the healthy gut microbiome is not as severely disturbed in these infections. Second, these patients have had little to no contact with health care leading up to their CDI and are more likely to be healthier than patients that have had extended hospital stays or repeated exposures to health care settings. This may also account for the differential expression of certain antibiotic resistance genes in patients with CDI. The microbiome of patients with hospital-acquired and healthcare-associated infections were more likely to express genes resistant to fosfomycin or bacitracin and sulfonamide, respectively. Research has shown that ambulatory and hospitalized patients harbor a high prevalence of antibiotic resistant organisms, regardless of whether they actually took antibiotics.[28]

Although there was a positive association between age and the risk for recurrence, the data from this study did not reveal any other associations between recurrence risk and other known factors for recurrence, including use of antibiotics and proton pump inhibitors.[2,29] Failure to find significant associations between these variables may be due to the small sample size, which was a major limitation to this study. Stool samples were obtained from the Connecticut Emerging Infections Program, which is responsible for sending positive *C. difficile* stool samples to the CDC as part of a nationwide surveillance program. The sample size was therefore limited by the number of positive samples that were available for metagenomic analysis. Due to the prospective study design, individuals were enrolled in the study before we knew whether or not they would experience a recurrent infection. This, in keeping with the estimate that approximately 20% of all positive *C. difficile* cases eventually experience a recurrent infection,

resulted in the number of recurrent cases in the study to be limited to only 7 samples. Consequently, many of the statistical analyses performed in this study were underpowered. Future studies that seek to explore the association between the gut microbiome and the risk for recurrent CDI should aim to increase the overall sample size.

Potential missing and incomplete medical histories for the subjects enrolled in the study was another major limitation. The majority of stool samples included in the analysis were from individuals who lived outside the New Haven catchment area. Three subjects' primary addresses were outside of Connecticut, while one subject moved out of state soon after their positive diagnosis. It is possible that individuals living outside of the New Haven catchment area received medical care from providers not in the Yale New Haven Health System (YNHHS). If so, it is not guaranteed that these medical visits were documented within the YNHHS Epic interface, and they may have been omitted during chart review. Consequently, information about patients' medical histories, including administered antibiotics and comorbidities, may have been incomplete.

This study offers a preliminary analysis of the association between the risk of recurrence of *Clostridioides difficile* infection and the human gut microbiome and resistome. Increasing our understanding of the factors that influence the risk of recurrence can aid in future efforts to identify individuals with CDI who are most likely to go on and experience a recurrent episode, as well as inform possible treatment strategies for patients for whom standard antibiotic therapies are ineffective in clearing infection. Whole genome sequencing provides unique insights into the relationship between CDI and the human microbiome and antibiotic resistome that is not provided by traditional sequencing methods. This pilot study may operate as a stepping stone for future metagenomic investigations into the human microbiome and the risk of CDI recurrence.

References

1. Lessa FC, Mu Y, Bamberg WM, et al. Burden of Clostridium difficile infection in the United States. 2015;372:825-34.

2. Fekety R, McFarland LV, Surawicz CM, Greenberg RN, Elmer GW, Mulligan ME. Recurrent Clostridium difficile diarrhea: characteristics of and risk factors for patients enrolled in a prospective, randomized, double-blinded trial. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 1997;24:324-33.

3. Zimlichman E, Henderson D, Tamir O, et al. Health care–associated infections: A meta-analysis of costs and financial impact on the US health care system. JAMA Internal Medicine 2013;173:2039-46.

4. Khanna S, Pardi DS, Aronson SL, et al. The epidemiology of community-acquired Clostridium difficile infection: a population-based study. Am J Gastroenterol 2012;107:89-95.

5. Peng Z, Jin D, Kim HB, et al. Update on antimicrobial resistance in Clostridium difficile: Resistance mechanisms and antimicrobial susceptibility testing. 2017;55:1998-2008.

6. Johanesen PA, Mackin KE, Hutton ML, et al. Disruption of the gut microbiome: Clostridium difficile infection and the threat of antibiotic resistance. Genes (Basel)2015:1347-60.

7. Lai KK, Melvin ZS, Menard MJ, Kotilainen HR, Baker S. Clostridium difficile-associated diarrhea epidemiology, risk factors, and infection control. Infection Control & Hospital Epidemiology 1997;18:628-35.

8. Pepin J, Saheb N, Coulombe MA, et al. Emergence of fluoroquinolones as the predominant risk factor for Clostridium difficile-associated diarrhea: a cohort study during an

epidemic in Quebec. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 2005;41:1254-60.

9. Pépin J, Valiquette L, Cossette B. Mortality attributable to nosocomial Clostridium difficile-associated disease during an epidemic caused by a hypervirulent strain in Quebec. CMAJ 2005;173:1037-42.

10. Dallal RM, Harbrecht BG, Boujoukas AJ, et al. Fulminant Clostridium difficile: An underappreciated and increasing cause of death and complications. Ann Surg 2002;235:363-72.

11. Hu MY, Katchar K, Kyne L, et al. Prospective derivation and validation of a clinical prediction rule for recurrent Clostridium difficile infection. Gastroenterology 2009;136:1206-14.

12. Schubert AM, Rogers MA, Ring C, et al. Microbiome data distinguish patients with Clostridium difficile infection and non-C. difficile-associated diarrhea from healthy controls. mBio 2014;5:e01021-14.

13. Zuo T, Wong SH, Lam K, et al. Bacteriophage transfer during faecal microbiota transplantation in Clostridium difficile infection is associated with treatment outcome. Gut 2018;67:634.

14. Milani C, Ticinesi A, Gerritsen J, et al. Gut microbiota composition and Clostridium difficile infection in hospitalized elderly individuals: a metagenomic study. Scientific Reports 2016;6:25945.

15. Rahman SF, Olm MR, Morowitz MJ, Banfield JF. Machine learning leveraging genomes from metagenomes identifies influential antibiotic resistance genes in the infant gut microbiome. mSystems 2018;3:e00123-17.

16. Khanna S, Montassier E, Schmidt B, et al. Gut microbiome predictors of treatment response and recurrence in primary Clostridium difficile infection. Alimentary Pharmacology & Therapeutics 2016;44:715-27.

17. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biology 2014;15:R46.

18. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science 2017;3:e104.

19. Yin X, Jiang XT, Chai B, et al. ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. Bioinformatics (Oxford, England) 2018;34:2263-70.

20. Pérez-Cobas AE, Artacho A, Ott SJ, Moya A, Gosalbes MJ, Latorre A. Structural and functional changes in the gut microbiota associated to Clostridium difficile infection. Front Microbiol 2014;5:335-.

21. Tvede M, Rask-Madsen J. Bacteriotherapy for chronic relapsing Clostridium difficile diarrhoea in six patients. The Lancet 1989;333:1156-60.

22. Lau SKP, Woo PCY, Woo GKS, et al. Eggerthella hongkongensis sp. nov. and eggerthella sinensis sp. nov., two novel Eggerthella species, account for half of the cases of Eggerthella bacteremia. Diagnostic Microbiology and Infectious Disease 2004;49:255-63.

23. Knecht H, Neulinger SC, Heinsen FA, et al. Effects of β-lactam antibiotics and fluoroquinolones on human gut microbiota in relation to Clostridium difficile associated diarrhea. PloS one 2014;9:e89417-e.

24. Wexler HM. Bacteroides: the Good, the Bad, and the Nitty-Gritty. Clinical Microbiology Reviews 2007;20:593.

25. Miquel S, Martin R, Rossi O, et al. Faecalibacterium prausnitzii and human intestinal health. Current opinion in microbiology 2013;16:255-61.

26. Yutin N, Makarova KS, Gussow AB, et al. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. Nature Microbiology 2018;3:38-46.

27. Draper LA, Ryan FJ, Smith MK, et al. Long-term colonisation with donor bacteriophages following successful faecal microbial transplantation. Microbiome 2018;6:220.

28. Levy SB, Marshall B, Schluederberg S, Rowse D, Davis J. High frequency of antimicrobial resistance in human fecal flora. Antimicrob Agents Chemother 1988;32:1801.

29. Kim JW, Lee KL, Jeong JB, et al. Proton pump inhibitors as a risk factor for recurrence of Clostridium-difficile-associated diarrhea. World J Gastroenterol 2010;16:3573-7.